THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 270, No. 20, Issue of May 19, pp. 12065–12074, 1995
Printed in U.S.A.

Yeast N-terminal Amidase

A NEW ENZYME AND COMPONENT OF THE N-END RULE PATHWAY*

(Received for publication, January 10, 1995, and in revised form, March 9, 1995)

Rohan T. Baker‡ and Alexander Varshavsky§

From the Division of Biology, California Institute of Technology, Pasadena, California 91125

The N-end rule relates the *in vivo* half-life of a protein to the identity of its N-terminal residue. Tertiary destabilizing N-terminal residues asparagine and glutamine function through their conversion, by enzymatic deamidation, into the secondary destabilizing residues aspartate and glutamate, whose activity requires their enzymatic conjugation to arginine, one of the primary destabilizing residues. We isolated a *Saccharomyces cerevisiae* gene, termed *NTA1*, that encodes an amidase (Nt-amidase) specific for N-terminal asparagine and glutamine. Alterations at the putative active-site cysteine of the 52-kDa Nt-amidase inactivate the enzyme. Null *nta1* mutants are viable but unable to degrade N-end rule substrates that bear N-terminal asparagine or glutamine. The effects of overexpressing Nt-amidase and other components of the N-end rule pathway suggest interactions between these components and the existence of a multienzyme targeting complex.

The *in vivo* half-lives of damaged or otherwise abnormal proteins are often shorter than half-lives of their normal counterparts. Many regulatory proteins are also metabolically unstable. Short lifetimes of regulatory proteins allow for rapid adjustments of their concentrations through changes in the rates of their synthesis or degradation. Features of proteins that confer metabolic instability are called degradation signals, or degrons (Varshavsky, 1992). The essential component of one degradation signal, termed the N-degron, is a destabilizing N-terminal residue of a protein (Bachmair *et al.*, 1986). A set of N-degrons containing different destabilizing residues is referred to as the N-end rule, which relates the *in vivo* half-life of a protein to the identity of its N-terminal residue (for review, see Varshavsky (1992)). The N-degron comprises two determinants: a destabilizing N-terminal residue and an internal lysine (or lysines) of a substrate (Bachmair and Varshavsky, 1989; Johnson *et al.*, 1990; Hill *et al.*, 1993; Dohmen *et al.*, 1994). The Lys residue is the site of formation of a multiubiquitin chain (Chau *et al.*, 1989; Arnason and Ellison, 1994; Spence *et al.*, 1995). Ubiquitin (Ub)[1] is a protein whose covalent conjugation to other proteins (often in the form of a multi-Ub chain) plays a role in a number of processes, primarily through routes that involve protein degradation (for review, see Hershko and Ciechanover (1992), Jentsch (1992), Varshavsky (1992), Hochstrasser (1992), and Ciechanover (1994)). Two classes of physiological N-end rule substrates were identified thus far. One is RNA polymerase of the Sindbis virus (and also, by inference, RNA polymerases of other alphaviruses) (deGroot *et al.*, 1991). The other is Gα subunit of heterotrimeric G proteins, a major class of signal transducers in eukaryotes (Madura and Varshavsky, 1994).

The N-end rule is organized hierarchically. In eukaryotes such as the yeast *Saccharomyces cerevisiae*, Asn and Gln are tertiary destabilizing N-terminal residues (denoted as N-d^t) in that they function through their conversion, by enzymatic deamidation, into the secondary destabilizing N-terminal residues Asp and Glu (denoted as N-d^s), whose activity requires their conjugation, by Arg-tRNA-protein transferase (R-transferase), to Arg, one of the primary destabilizing N-terminal residues (denoted as N-d^p) (Gonda *et al.*, 1989; Balzi *et al.*, 1990). The N-d^p residues are bound directly by *N*-recognin (also called E3), the recognition component of the N-end rule pathway (Varshavsky, 1992). In *S. cerevisiae*, *N*-recognin is a 225-kDa protein (encoded by *UBR1*) that selects potential N-end rule substrates by binding to their N-d^p residues Phe, Leu, Trp, Tyr, Ile, Arg, Lys, or His (Bartel *et al.*, 1990; Baker and Varshavsky, 1991; Madura *et al.*, 1993).

An apparently enzymatic conversion of N-terminal Asn and Gln in cytosolic proteins into Asp and Glu was demonstrated in both yeast and mammalian cells (Gonda *et al.*, 1989), but the postulated amidase(s) responsible for this conversion remained unknown. Stewart *et al.* (1994, 1995) described purification, cDNA isolation, and analysis of porcine N-terminal amidase (Nt-amidase) that deamidates N-terminal Asn but not Gln. We report a *S. cerevisiae* gene *NTA1* that encodes an enzymatically distinct Nt-amidase; it deamidates either N-terminal Asn or Gln and is essential for the *in vivo* degradation of proteins bearing N-terminal Asn or Gln residues. (Nta1p has previously been referred to as Dea1p (Varshavsky, 1992).) In addition, we examine functional interactions between targeting components of the N-end rule pathway and consider a multienzyme targeting complex whose components include Nt-amidase (Nta1p), R-transferase (Ate1p), *N*-recognin (Ubr1p), and Ubc2p, one of at least 11 Ub-conjugating enzymes in *S. cerevisiae*.

EXPERIMENTAL PROCEDURES

Strains, Genetic Techniques, and β-Galactosidase Assay—*S. cerevisiae* strains were BWG1 7a (*MAT***a** *his4 ura3 ade1 leu2*) (Guarente *et*

* This work was supported by Grants DK39520 and GM31530 from the National Institutes of Health (to A. V.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) L35564.

‡ Supported by fellowships from the Fulbright Foundation and the Life Sciences Research Foundation. Present address: Molecular Genetics Group, John Curtin School of Medical Research, Australian National University, P. O. Box 334, Canberra, ACT 2601, Australia.

§ To whom correspondence should be addressed: Div. of Biology, 147-75, Caltech, Pasadena, CA 91125. Tel.: 818-395-3785; Fax: 818-440-9821; E-mail: varshavskya@starbase1. caltech.edu.

[1] The abbreviations used are: Ub, ubiquitin; N-d^t, a tertiary destabilizing N-terminal residue; N-d^s, a secondary destabilizing N-terminal residue; N-d^p, a primary destabilizing N-terminal residue; Nt-amidase, amidase specific for N-terminal Asn and Gln; R-transferase, Arg-tRNA-protein transferase; βgal, *E. coli* β-galactosidase; X-gal, 5-bromo-4-chloro-3-indolyl β-D-galactoside; wt, wild-type; kb, kilobase(s); ORF, open reading frame; bp, base pair(s); PAGE, polyacrylamide gel electrophoresis.

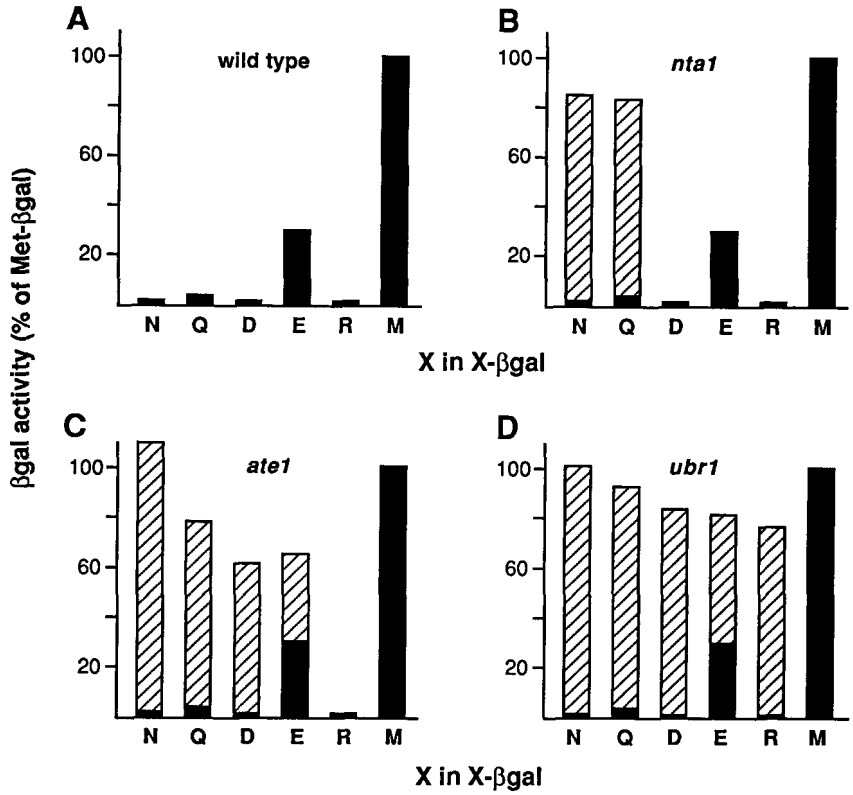


FIG. 1. **Three classes of mutants that stabilize short-lived Asn-βgal.** *A*, levels of βgal activity (plotted as percentages of Met-βgal activity) in extracts from the parental (wt) *S. cerevisiae* strain BWG1–7a expressing different X-βgal (Ub-*X*-βgal) test proteins. *B*, *C*, and *D*, representative mutants RBY56 (*nta1–1*), RBY42 (*ate1*), and RBY45 (*ubr1*), respectively. These mutants were isolated through their inability to degrade the normally short-lived Asn-βgal. Each of the strains was transformed with plasmids expressing X-βgals whose N termini bore either tertiary (N, Asn; Q, Gln) or secondary (D, Asp; E, Glu) destabilizing residues and also a primary destabilizing (R, Arg) or a stabilizing (M, Met) residue. The *striped* upper portions of the *bars* indicate differences between the activities of a given X-βgal in the wt strain (*A*) and in its mutant derivatives (*B–D*). Values shown are the means from duplicate measurements, which yielded results within 10% of each other.

al., 1982), BWG9a-1 (*MATα his4 ura3 ade6*) (Bartel *et al.*, 1990), RBY56 (*MAT***a** *nta1–1 his4 ura3 ade1 leu2*); RBY561 (*MATα nta1–1 his4 ura3 ade6 leu2*), and YBY1 (*MAT***a** *nta1-Δ1::URA3 his4 ura3 ade1 leu2*) (this work). Cells were grown in rich (YPD) or synthetic media (Sherman *et al.*, 1986), with the latter containing either 2% glucose (SD medium) or 2% galactose (SG medium). X-gal plates also contained 0.1 M K_2HPO_4-KH_2PO_4 (pH 7.0) and 5-bromo-4-chloro-3-indolyl β-D-galactoside (X-gal) at 40 μg/ml (Guthrie and Fink, 1991). *Escherichia coli* strains MC1061, JM101, and DH5αF′ (Ausubel *et al.*, 1992) were grown in Luria Broth (LB) or in M9 minimal medium containing auxotrophic nutrients and antibiotics as required. Transformation of *S. cerevisiae* was carried out using the lithium acetate method (Ito *et al.*, 1983; Baker, 1991). Yeast mating, sporulation, and tetrad analyses were performed as described previously (Guthrie and Fink, 1991). *S. cerevisiae* were cured of *URA3*-expressing plasmids using 5-fluoroorotic acid (Boeke *et al.*, 1984). Activity of βgal in yeast extracts was measured using *o*-nitrophenyl β-D-galactoside (Baker and Varshavsky, 1991).

Isolation of nta1 Mutants—BWG1–7a was transformed with pUB23-N, a high copy (2 μm-based), *URA3*-containing plasmid expressing Ub-Asn-βgal from the galactose-inducible $P_{CYC1/GAL10}$ hybrid promoter (Bachmair *et al.*, 1986). The cells were mutagenized with ethyl methanesulfonate (Sherman *et al.*, 1986) to ~20% survival and plated on SD medium lacking uracil. After ~3 days of growth at 23 °C, colonies were replica plated onto X-gal-SG plates and incubated for 2 days at 23 or 36 °C. About 100 colonies (out of ~3×10^4 colonies screened) that turned blue at 23 °C and ~30 colonies that turned blue only at 36 °C were picked and retested. Extracts from 43 colonies (23 °C) and 15 colonies (36 °C) that passed the second X-gal test were examined for βgal activity using the *o*-nitrophenyl β-D-galactoside assay. Thirty putative 23 °C mutants and eight putative temperature-sensitive (*ts*) mutants with high βgal activity were cured of pUB23-N using 5-fluoroorotic acid and were retransformed with the same (unmutagenized) plasmid. Seven apparently plasmid-linked mutants were eliminated at this stage. The rest were transformed with pUB23-X plasmids expressing Ub-Met-βgal, Ub-Pro-βgal, Ub-Asn-βgal, Ub-Asp-βgal, or Ub-Arg-βgal (Bartel *et al.*, 1990). This analysis segregated each mutant into one of three classes: (i) those with increased levels of Asn-βgal, Asp-βgal, and Arg-βgal (23 mutants); (ii) those with increased levels of Asn-βgal and Asp-βgal (two mutants); and (iii) those with increased levels of Asn-βgal but not Asp-βgal or Arg-βgal (six mutants). All mutants had wt levels of Met-βgal and Ub-Pro-βgal. Mutants of the third class were examined by complementation and found to harbor recessive mutations in a single complementation group. One mutant, RBY56, was crossed to the parental strain BWG9a-1; sporulation and tetrad analysis of the resulting diploid showed a 2:2 segregation of high Asn-βgal levels. Two back-crosses of RBY56 to BWG9a-1 and BWG1–7a yielded the haploid RBY561 carrying the original mutation. Curing RBY561 of pUB23-N, transforming it with pUB23-X plasmids expressing different Ub-*X*-βgals (Bachmair and Varshavsky, 1989; *X* = Asn, Gln, Asp, Glu, Leu, and Arg), and measuring βgal activity in the transformants (Fig. 1*B*) showed that among the normally short-lived βgals, only those bearing N-d^c residues (Asn and Gln) accumulated to high levels in RBY561. Pulse-chase analyses of X-βgals (Fig. 2) confirmed these results. The gene a mutation in which produced the phenotype of RBY561 was termed *NTA1*.

Isolation of NTA1—The *nta1–1* strain RBY561, carrying pNL, a derivative of pUB23-N (expressing Ub-Asn-βgal) in which the *URA3* marker had been replaced by *LEU2*, was transformed with *S. cerevisiae* genomic DNA library carried in the *URA3*, *CEN4*-based vector YCp50 (Rose *et al.*, 1987). About 4×10^3 transformants were screened for white colonies (low levels of βgal) on SG-X-gal plates that lacked uracil and leucine. Two of the six initially selected transformants remained white upon retesting on SG-X-gal plates; these results were confirmed using the *o*-nitrophenyl β-D-galactoside assay for βgal. When the two transformants were cured of their library-derived plasmids on 5-fluoroorotic acid plates and retested on SG-X-gal plates, only one isolate yielded blue colonies (high levels of βgal). Plasmid DNA (Hoffman and Winston, 1987) from these cells was used to transform *E. coli* MC1061 to ampicillin resistance. Transformants carrying YCp50 library-derived plasmids were distinguished from those carrying pNL by picking white *E. coli* colonies on LB/ampicillin/X-gal plates. A ~4.3-kb subclone of the ~14.3-kb insert in the plasmid thus obtained (termed pRB8) complemented the defect in degradation of Asn-βgal in RBY561 and the other five *nta1* mutants. This subclone was sequenced, revealing two complete open reading frames (ORFs) and a portion of a third (Fig. 3*A*). Comparisons of the amino acid sequences encoded by these ORFs with sequences in GenBank showed that one of the two complete ORFs had similarities to known amidotransferases. A low copy plasmid expressing only the putative *NTA1* portion of the ~4.3-kb insert (Fig. 3*A*) was constructed and found to complement all six *nta1* mutants. Conversely, the low copy plasmid pRB8E2.5 that expressed only the smaller complete ORF was unable to complement *nta1* mutants.

The nta1-Δ1 Allele—To construct *nta1-Δ1*, the ~3.7-kb *Xho*I-*Sal*I fragment of pRB8S13 (Fig. 3*A*) that contained *NTA1* was subcloned into *Sal*I-cut pUC9, yielding pUC3.7. The yeast *URA3* gene, isolated as a 1.1-kb *Xba*I-*Kpn*I fragment of pRBU1 (Tobias and Varshavsky, 1991),

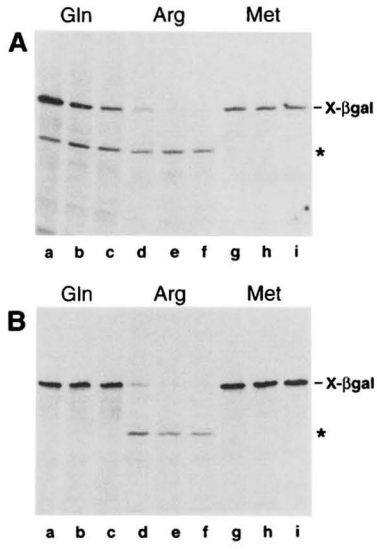


FIG. 2. **Metabolic stabilization of Gln-βgal in the *nta1–1* mutant.** *A*, *lanes a–c*, BWG1–7a (wt) cells that expressed Gln-βgal (Ub-Gln-βgal) were labeled with Tran^{35}S-label for 5 min at 30 °C, followed by a chase for 0, 10, and 30 min, respectively; extraction and immunoprecipitation of test proteins with an antibody to βgal; and SDS-PAGE and fluorography (see "Experimental Procedures"). *Lanes d–f*, same as *lanes a–c* but with Arg-βgal; *lanes g–i*, same as *lanes a–c* but with Met-βgal. *B*, same as the corresponding lanes in *A* but with RBY561 (*nta1–1*) cells. The X-βgal bands are indicated. An asterisk denotes a ~90 kDa, long-lived βgal cleavage product characteristic of short-lived βgal-based proteins (Bachmair *et al.*, 1986; Baker and Varshavsky, 1991). Note the absence of this species from *nta1–1* cells expressing Gln-βgal. Note also the much weaker labeling of Arg-βgal (in comparison with longer-lived X-βgals) even at the beginning of chase; this "zero-point effect" (Baker and Varshavsky, 1991) is caused by degradation of Arg-βgal during the labeling.

was ligated between the *Spe*I site at the 5′ end of the *NTA1* ORF and the *Kpn*I site near the 3′ end of the *NTA1* ORF in pUC3.7, yielding p3.7::URA3, which lacked 82% of the *NTA1* ORF (Fig. 3*A*). The ~3.5-kb *Hin*dIII fragment of p3.7::URA3 containing *nta1-Δ1::URA3* was used to replace *NTA1* in *ura3* strains of *S. cerevisiae* by homologous recombination (Rothstein, 1991). Southern hybridization analysis of Ura$^+$ transformants, using restriction endonuclease cuts diagnostic of the transplacement, was used to confirm the predicted structure of the integrated nta1-Δ1 allele (data not shown).

To verify that the isolated NTA1 gene was allelic to the gene of the original *nta1–1* mutation, a diploid strain YRB200 was constructed by crossing RBY561 (*nta1–1*) to YRB1 (*nta1-Δ1::URA3*). This strain was unable to degrade N-end rule substrates bearing N-terminal Asn or Gln, but it retained the ability to degrade N-end rule substrates of other classes (data not shown). Since both the *nta1–1* and *nta1-Δ1::URA3* mutations were recessive, this result indicated that the cloned *NTA1* gene was allelic to *nta1–1*.

Other DNA Constructs—DNA fragments were isolated from agarose gels using Geneclean (Bio 101). To verify that expression of the *NTA1* ORF was sufficient for complementation of the *nta1* phenotype, a low copy plasmid was constructed whose only yeast ORF was *NTA1*. The 944-bp *Sca*I-*Eco*RI fragment of pRB8S13 that contained the 5′ region of *NTA1* was ligated into *Sma*I/*Eco*RI-cut pRS316 (a *CEN4*, *URA3*-based plasmid; Sikorski and Hieter (1989)), yielding p316ESc. The 998-bp *Eco*RI fragment of pRB8S13 that contained the rest of *NTA1* (Fig. 3*A*) was then ligated into the *Eco*RI site of p316ESc, yielding pNTA1, which contained the reassembled *NTA1* ORF, in addition to 335 bp of yeast DNA 5′ to the (inferred) start codon of *NTA1* and 236 bp of yeast DNA 3′ to *NTA1* stop codon. An *NTA1*-overexpressing plasmid was constructed by ligating the ~2-kb *Bam*HI-*Sal*I fragment of pNTA1 into *Bam*HI/*Sal*I-cut YEplac195 (a 2 μm, *URA3*-based vector; Gietz and Sugino (1988)), yielding p195NTA1. To construct the *NTA1-ha* allele, the ~2-kb *Bam*HI-*Sal*I fragment of pNTA1 was ligated into *Bam*HI/*Sal*I-cut RF DNA of the phage M13mp18 (Ausubel *et al.*, 1992). A synthetic oligodeoxynucleotide (5′-GTATAGCTTCTGTTCATCATC-CAAGCTAGCGTAATCTGGAACATCGTATGGGTAATCATCTTCAGG-ATATCC-3′) was used for insertional mutagenesis with the MutaGene kit (Bio-Rad), yielding an ORF that encoded a modified Nta1p (Nta1p-ha) containing the sequence Tyr-Pro-Tyr-Asp-Val-Pro-Asp-Tyr-Ala-Ser-Leu-Asp between residues 112 and 113 of Nta1p. This sequence contained the ha epitope recognizable by the monoclonal antibody 12CA5 (Johnson *et al.*, 1992). The insertion (containing a diagnostic *Nhe*I site) was verified by *Nhe*I mapping and DNA sequencing, and the ~1-kb *Bam*HI-*Sal*I fragment of the resulting plasmid was ligated into *Bam*HI/*Sal*I-cut YEplac195, yielding p195NTA1-ha. To mutate the Cys-187 residue of Nta1p, two oligonucleotides (5′-CTTAAGTCCATAGATATC-CCTATTGATGC 3′ and 5′-GGACTTAAGTCCATCGCGATACCTATT-GATGC-3′) were used with the M13mp18 subclone described above to produce ORFs encoding Nta1p-C187S and Nta1p-C187A, in which Cys-187 was converted, respectively, into Ser and Ala. The oligos contained diagnostic restriction sites (*Eco*RV and *Nru*I, respectively); both mutations were confirmed by sequencing as well. The ~2-kb *Bam*HI-*Sal*I fragments of the resulting plasmids were ligated into *Bam*HI/*Sal*I-cut YEplac195, yielding p195NTA1-CS and p195NTA1-CA, which expressed, respectively, Nta1p-C187S and Nta1p-C187A from the P_{NTA1} promoter. The ha tag was linked to Nta1p-C187S and Nta1p-C187A by ligating the 472-bp *Spe*I-*Hin*dIII fragment of p195NTA1-ha, together with either the 897-bp *Hin*dIII-*Xba*I fragment of p195NTA1-CS or an analogous fragment of p195NTA1-CA, into *Spe*I/*Xba*I-cut p195NTA1-ha, yielding, respectively, p195NTA1-CS-ha and p195NTA1-CA-ha.

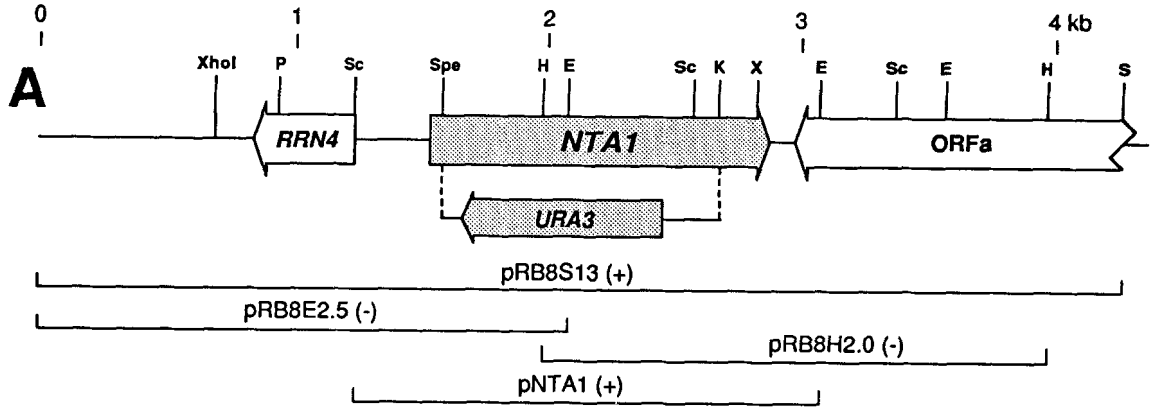
Plasmids expressing *S. cerevisiae* Ate1p either alone or together with Nta1p were constructed as follows. The ~4-kb *Hin*dIII fragment of pA4-H4 (Balzi *et al.*, 1990) that contained *ATE1* was subcloned into the *Hin*dIII-cut high copy plasmid YEplac195, yielding p195ATE1. The same fragment was also subcloned into the low copy plasmid YCplac33 (Gietz and Sugino, 1988), yielding p33ATE1. pNTA1/ATE1, a high copy plasmid that overexpressed both Nta1p and Ate1p, was constructed by subcloning the above *Hin*dIII fragment into p195NTA1 that had been partially digested with *Hin*dIII, and screening for products in which the insert was located at the polylinker *Hin*dIII site.

Southern Hybridization and DNA Sequencing—DNA was isolated from *S. cerevisiae* as described by Hoffman and Winston (1987), digested with restriction endonucleases, and processed for electrophoresis and hybridization as described by Bartel *et al.* (1990). Restriction fragments of the *NTA1*-containing a 4.3-kb fragment of pRB8S13 (Fig. 3*A*) were subcloned into M13mp18 and M13mp19 (Ausubel *et al.*, 1992) and were sequenced using the Sequenase kit (U. S. Biochemical Corp.). The entire 4.3-kb fragment was sequenced on both strands. The nucleotide sequence of *NTA1* (GenBank accession L35564) and the predicted amino acid sequence of Nta1p (Fig. 3*B*) were compared with sequences in the GenBank®/EMBL data base using the GCG program (Deveraux *et al.*, 1984) (version 7.2, Genetics Computer Group, Madison, WI).

5′ Mapping of NTA1 mRNAs—For mapping by primer extension, RNA was isolated from an exponential culture of BWG1–7a in YPD, using the phenol/chloroform/glass bead procedure of Sprague *et al.* (1983). Poly(A)$^+$ RNA was isolated from total RNA by chromatography on oligo(dT)-agarose (Pharmacia Biotech Inc.). Primer extension analysis was carried out as described by Teem and Rosbash (1983), using 3 μg of poly(A)$^+$ RNA/sample. The oligodeoxynucleotide primer was complementary to the coding strand of *NTA1* between nucleotide positions +34 and +67. For mapping by S1 nuclease, the procedure described by Nyunoya and Lusty (1984) was used, with 50 μg of total RNA or 2 μg of poly(A)$^+$ RNA/sample. The probe was a single-stranded DNA from an M13 subclone containing the 400-bp *Sca*I-*Spe*I fragment of pRB8S13 (*NTA1* nucleotide positions −338 to +63) uniformly labeled with ^{32}P (Baker *et al.*, 1992).

Phenotypic Characterization of the nta1-Δ1 Mutant—Assays measuring sensitivity of yeast cells to chronic heat stress (at 39 °C), sensitivity to canavanine, and survival in stationary phase were carried out as described by Finley *et al.* (1987). Sensitivity to acute heat stress was determined by exposing cells (which have been growing exponentially in YPD at 30 °C) to YPD at 50 °C for 0–15 min prior to plating on YPD to assay colony formation at 30 °C. Ability to use glycerol as a carbon source was assayed on YPD plates lacking glucose and containing 3% (v/v) glycerol. Ability to utilize Asn or Gln as a nitrogen source was assayed on synthetic media plates containing 0.17% yeast nitrogen base (without amino acids and ammonium sulfate; Difco), 2% glucose, auxotrophic nutrients at standard concentrations (Sherman *et al.*, 1986), and 0.1% (w/v) of either Asn or Gln as a major nitrogen source. Control plates lacked Asn and Gln, and in addition either contained or lacked 0.1% (w/v) ammonium sulfate.

We did not detect a significant difference between the *S. cerevisiae* *nta1-Δ1* mutant and a congenic wt strain in their sensitivity to acute or chronic heat stress; in their survival at stationary phase after growth in rich or poor media, or upon starvation for either carbon or nitrogen; in their sensitivity to canavanine (a toxic arginine analog); in their ability



B

```
MLIDAIHGAKMSTKLLVSLKVLVIQLNPQIGQVDQTIKRTWSILDKVTKS      50
ATYVKPDIILFPEFALTGYSFHARKDILPYVTKKDEGPSFELAKSISEKF     100
QCYTIIGYPEDDDEQKLYNSALVVNPQGEQIFNYRKTFLYDTEMNWDCEE    150
NPEGFQTFPMDFSKCAKLSNEDSYNRDVTLKASIGICMDLSPYKFMAPFN    200
HFEFSSFCVDNNVELILCPMAWLNSTSITDKQTLHNNSLLEAAKNKIAFA    250
LKEQGLPLAGSQGIYQLKIGDSQRTPRVPSDDSTSEYKDMDEPDMSNVNY   300
WILRFFPFLYFKSRINWFKNSSLIESILGKTRMPLDHEYYKDGKHKEDTI    350
DLLDSEEVIKDTVLEKTFLGTSLGQPWKFQGKNAILVLANRCGTEDGTTI   400
FAGSSGIYKFNGKKPKGSQDDDESSLDSLNESVELLGNLGKGLEGAILRE    450
VQFEVFR                                                457
```

C

```
                    I  F  V  L
consensus:      L P I I G I C M G H Q
Nta1 (181-191): K A S I G I C M D L S
```

D

```
-90
AATAAAACCACTTTAGTTTCACACATTTTAAGCATTTCAATGCTAAAAAGTGCTAGTTTACAGCGA

-24   o  ●●o      oo  ooooo      o          +31
TAACTCTATCGTGACATTCAGTGAATGCTAATAGACGCAATTCATGGTGCTAAGATGAGCACAAAACTT...
                    M  L  I  D  A  I  H  G  A  K  M  S  T  K  L...
```

FIG. 3. **The *NTA1* locus.** *A*, ORFs are shown as *arrow-shaped boxes* indicating the direction of transcription. The incompletely sequenced ORFa is shown as a *jagged-end box*. Subcloned regions are also indicated, with (+) or (−) denoting their ability to complement the *nta1–1* phenotype. *Dashed lines* delineate the region of *NTA1* that has been replaced with *URA3* in the *nta1-Δ1* allele. Restriction sites: *E*, *Eco*RI; *H*, *Hin*dIII; *K*, *Kpn*I; *P*, *Pst*I; *S*, *Sal*I; *Sc*, *Sca*I; *Spe*, *Spe*I; *X*, *Xba*I. The scale (in kb) is *above* the map, with *zero* denoting the end of yeast genomic DNA insert in pRB8. Nucleotide sequence encompassing *NTA1* (a 3747-bp region from the *Xho*I to the *Sal*I site) has been submitted to GenBank (accession number L35564). An ORF located 346 bp upstream of *NTA1* and oriented in the opposite direction has been identified as the *RRN4* gene encoding the 125-residue A12.2 subunit of *S. cerevisiae* RNA polymerase I (Nogi *et al.*, 1993). A partially sequenced ORFa 156 bp downstream of *NTA1* is oriented in the opposite direction and encodes a protein of at least 740 residues. ORFa is transcriptionally active (data not shown) and encodes a protein highly similar to the product of a putative *S. cerevisiae* ORF on chromosome XI (GenBank accession numbers Z28200 and Z28201). The functions of either of these ORFs are unknown. *B*, deduced amino acid sequence of the Nta1p. Amino acid residues are numbered on the *right*. The sequence Ile-Gly-Ile-Cys-Met that matches a portion of the consensus active-site region of several amidotransferases and contains an essential Cys residue is *doubly underlined*. A *black rectangle* indicates the position of the 12-residue ha epitope tag in the Nta1p-ha derivative of Nta1. *C*, alignment of the Nta1p sequence with the 11-residue consensus sequence encompassing the putative active-site Cys residue of glutamine amidotransferases (Nyunoya and Lusty, 1984). Alternate amino acids in the consensus among these enzymes are shown above the consensus sequence. The region of identity between Nta1p and the consensus sequence is *boxed*. The essential Cys-187 of Nta1p is *doubly underlined*. *D*, nucleotide sequence at the 5′ region of *NTA1*. Positions of the major and minor 5′ ends of *NTA1* mRNAs are indicated by *closed* and *open circles*, respectively. ATG codons are *boxed*. A start codon at position +1 was inferred in part from the 5′ mapping data. The in-frame ATG codon at position +31 is likely to be used as an *in vivo* translation initiation site as well (see text). The ATG codon at position −51 is indicated by an *asterisk*. This in-frame ATG is absent from the presently detectable *NTA1* mRNAs; it is followed by two in-frame (*underlined*) stop codons. Motifs that are present in the promoter regions of genes encoding components of the N-end rule pathway (Fig. 6) are *doubly underlined*.

to grow on glycerol as a carbon source; and in their ability to utilize either Asn or Gln as a source of nitrogen. No short-lived yeast proteins detectable by a pulse-chase and two-dimensional electrophoresis were significantly stabilized in the *nta1-Δ1* mutant (data not shown). *ubr1Δ* mutants (in which normally short-lived N-end rule substrates are metabolically stable) grow slightly (~3%) slower than wt cells and have a defect in sporulation, increased fraction of asci with fewer than four spores (Bartel *et al.*, 1990). The growth rate phenotype was not observed with the *nta1-Δ1* mutant (data not shown), while the sporulation of an *nta1Δ/nta1Δ* strain has yet to be investigated.

Protein Labeling, Pulse-Chase Analysis, and Immunoblotting—Pulse labeling with Tran^{35}S-label (ICN), a chase in the presence of cycloheximide, preparation of cell extracts, immunoprecipitation with a polyclonal antibody to βgal (Sigma), and electrophoretic analysis of X-βgals by SDS-PAGE in 6% gels were carried out as described by Bachmair *et al.* (1986), with slight modifications (Baker and Varshavsky, 1991). Immunoblotting of extracts from the *nta1–1* strain RBY561 that has been transformed with either p195NTA1, p195NTA1-ha, p195NTA1-ha-CS, or p195NTA1-ha-CA was carried out after SDS-PAGE in a 10% gel, using a monoclonal anti-ha antibody (Bartel *et al.*, 1990; Madura *et al.*, 1993), a phosphatase-linked second antibody, a chromogenic phosphatase substrate, and procedures described by Tobias and Varshavsky (1991) and Baker *et al.* (1992).

RESULTS AND DISCUSSION

Isolation of nta1 Mutants and Cloning of the NTA1 Gene—To screen for *S. cerevisiae* mutants defective in the N-end rule pathway, we used a strain carrying a plasmid that expressed Ub-Asn-βgal. Ub fusions are rapidly cleaved *in vivo* after the last residue of Ub, making possible the production of otherwise identical proteins bearing different N-terminal residues (Bachmair *et al.*, 1986; Baker *et al.*, 1992). Since Asn-βgal is short-lived in wt cells ($t_{1/2}$ of ~3 min at 30 °C; Bachmair and Varshavsky (1989)), its steady-state level is low, and the corresponding yeast colonies are white on plates containing the chromogenic βgal substrate X-gal (Bartel *et al.*, 1990). By contrast, cells that express long-lived X-βgals such as Met-βgal ($t_{1/2}$ > 30 h) have high βgal activity and form blue colonies on X-gal plates. Cells expressing Asn-βgal were mutagenized, plated on X-gal plates, and screened for blue colonies. These were tested further, and the putative *nta1* mutants among them were identified as described under "Experimental Procedures."

The *NTA1* gene was cloned by complementation (see "Experimental Procedures"). The position of the start (ATG) codon of the *NTA1* ORF was inferred so as to yield the largest ORF (Fig. 3). The 1,371-bp *NTA1* encodes an acidic (calculated pI of 4.9), 457-residue (51.8 kDa) protein. The codon adaptation index of *NTA1* (calculated according to Sharp and Li (1987)) is 0.125, characteristic of weakly expressed yeast genes.

An Essential Cysteine in Nta1p—Weak sequence similarities were detected between Nta1p and an aliphatic amidase from *Pseudomonas aeruginosa* (Ambler *et al.*, 1987) as well as several other amidotransferases. The substrates of aliphatic amidase (acetamide and propionamide) are the side chains of Asn and Gln, which are the substrates of Nta1p when these residues are present at the proteins' N termini. Nyunoya and Lusty (1986) identified an 11-residue region conserved among 7 glutamine amidotransferases from five species, including *E. coli*, *S. cerevisiae*, and *Neurospora crassa* (Fig. 3*C*). The conserved region contains a Cys residue that could be labeled with reactive glutamine analogs in two of these enzymes, suggesting that this cysteine is a part of the active center (Nyunoya and Lusty, 1986). Nta1p contains a 5-residue sequence that is identical to the sequence in the middle of the 11-residue consensus stretch and includes the conserved cysteine (Cys-187) (Fig. 3*C*).

We converted Cys-187 of Nta1p into Ser and Ala. The resulting Nta1p-C187S and Nta1p-C187A (expressed from either low or high copy plasmids) lacked Nt-amidase activity, as inferred from their inability to restore the degradation of Asn-βgal in the *nta1–1* mutant (Fig. 4 and data not shown). Nta1p was then

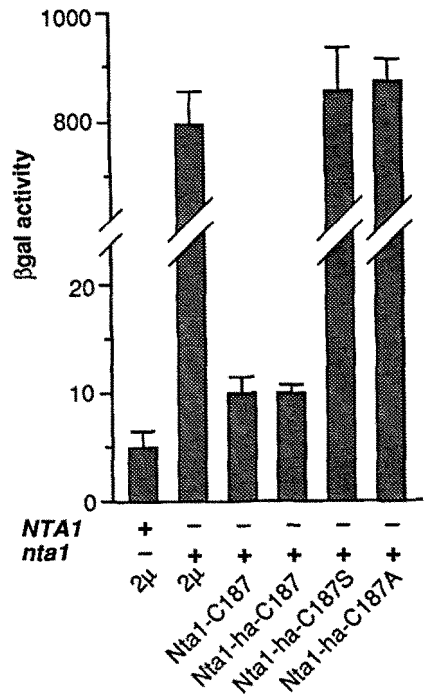


FIG. 4. **Nt-amidase contains an essential cysteine residue.** Substitution of Cys-187 in Nta1p with either Ser or Ala results in metabolic stabilization of Asn-βgal. Congenic *S. cerevisiae* strains BWG1–7a (*NTA1*) and RBY561 (*nta1–1*) carrying a plasmid that expressed Asn-βgal were transformed with either a high copy vector YEplac195 (a control) or an otherwise identical plasmid expressing either the wt Nta1p, the ha epitope-tagged Nta1p (Nta1p-ha), or Nta1p-ha in which Cys-187 was converted into either Ser (Nta1p-ha-C187S) or Ala (Nta1p-ha-C187A). Extracts from cultures in exponential growth were assayed for βgal activity. Values shown are the means of at least three independent measurements. Standard deviations are indicated *above* the *bars*. Note a discontinuity in the ordinate scale.

tagged with a 12-residue sequence containing the ha epitope, making it possible to immunoprecipitate Nta1p-ha with an anti-ha antibody (Field *et. al.*, 1988; Johnson *et al.*, 1992). The ha tag was positioned within a putative loop between residues 112 and 113 of Nta1p (Fig. 3*B*). Nta1p-ha and the unmodified Nta1p were equally effective in restoring the degradation of Asn-βgal in the *nta1–1* mutant, whereas no complementation was observed with Nta1p-ha-C187S or Nta1p-ha-C187A (Fig. 4). Internal tagging of Nta1p was necessitated by inactivity of the C-terminally tagged Nta1p (data not shown) and by an *a priori* drawback of N-terminal tagging, given uncertainties about the location of a start codon in *NTA1* (see below). The levels of Nta1p-ha in cells, measured by immunoblotting of cell extracts after SDS-PAGE, were the same irrespective of whether Nta1p-ha contained or lacked a substitution at Cys-187 (Fig. 5). Thus, Cys-187 is required for the activity of Nta1p, in agreement with the observation that a thiol-blocking reagent *N*-ethylmaleimide inhibits the conversion of N-terminal Gln into Glu in reticulocyte extract (Gonda *et al.*, 1989).

Location of Start Codons in NTA1—There is an ATG codon 51 bp upstream of (and in frame with) the inferred *NTA1* start codon but with two stop codons in between (Fig. 3*D*). If *NTA1* mRNA were to contain the −51 ATG codon, initiation of translation at this codon would result in the synthesis of an 8-residue peptide and might also interfere with initiation at the downstream (+1) ATG codon. A short translated ORF is present in the 5′ leader region of the yeast *GCN4* mRNA, which encodes transcriptional activator of the regulon for amino acid biosynthesis (Hinnebusch and Liebman, 1991). A potentially more relevant example is *CPA1*, which encodes glutamine

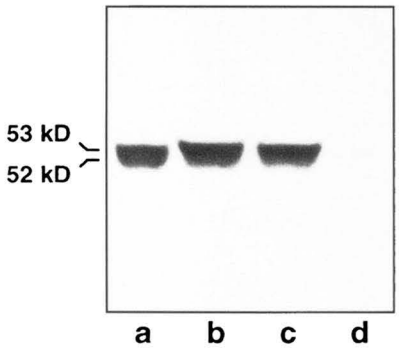


FIG. 5. **The Nta1p protein.** Immunoblotting of extracts from cells that expressed Nta1p-ha containing either the wt Cys-187 (Nta1p-ha), Ser-187 (Nta1p-ha-CS), or Ala-187 (Nta1p-ha-CA). Equal amounts of total protein in extracts from the *nta1–1* strain RBY561 that has been transformed with plasmids expressing Nta1p-ha-CA (*lane a*), Nta1p-ha-CS (*lane b*), Nta1p-ha (*lane c*), or the untagged (*control*) Nta1p (*lane d*) were fractionated by SDS-PAGE and analyzed by immunoblotting with anti-ha antibody (see "Experimental Procedures"). Note the presence of two Nta-ha species (~53 kDa and ~52 kDa; see text).

amidotransferase, a subunit of carbamoyl-phosphate synthetase and component of the arginine biosynthetic pathway. *In vivo* translation of the upstream ORF in *CPA1* mRNA yields a 25-residue peptide that down-regulates translation of the major *CPA1* ORF in the presence of arginine (Werner *et al.*, 1987).

To determine whether *NTA1* mRNA contains the −51 start codon, we used primer extension and S1 nuclease mapping (see "Experimental Procedures"). Both tests identified major 5′ ends of *NTA1* mRNA at positions −14 and −15 relative to the inferred (+1) start codon, with less abundant 5′ ends at −15 and −13 (Fig. 3*D*). No transcripts extending beyond −19 were detected, indicating that most of *NTA1* mRNAs lack the (−51) start codon. Thus, *NTA1* appears not to be regulated in a way observed with *CPA1*. The 5′ mapping also showed that in a minor but significant fraction of *NTA1* mRNAs, their 5′ ends are located immediately upstream, or even downstream of the inferred (+1) *NTA1* start codon. Specifically, S1 mapping detected minor 5′ ends largely at positions −4, −3, and +1, while primer extension detected sites at +2, +3, +4, +5, and also at +13 relative to the inferred (+1) start codon (the 5′ end at +13 was not detected by S1 mapping) (Fig. 3*D* and data not shown). Thus, there exist *NTA1* mRNAs that either lack the inferred (+1) start codon or contain it too close to the 5′ end of the message for efficient initiation of translation at that position. An in-frame ATG is present 30 bp downstream of the inferred (+1) *NTA1* start codon (Fig. 3*D*). Initiation of translation at this (+31) ATG should yield a 50.8-kDa protein lacking the first 10 residues of the inferred Nta1p (51.8 kDa). The (+31) ATG lies within a relatively favorable context for translation initiation (Kozak, 1992), with As in positions −3 and +4 (AAGATGA), whereas the (+1) ATG is located in a less favorable context, with pyrimidines at −3 and +4 (TGAATGC).

Immunoblot Analysis of Nta1p-ha—Two nearly comigrating, Nta1p-specific, ha-containing species of ~52 and ~53 kDa, were observed upon immunoblot analysis of Nta1-ha, the smaller protein being less abundant and partially obscured by the band of the ~53-kDa Nta1p-ha (Fig. 5 and data not shown). The ~1-kDa difference between these species of Nta1p is consistent with the possibility that the translation start site of the larger (~53 kDa) Nta1p-ha is at the inferred (+1) ATG codon of *NTA1* (predicted Nta1p-ha of 53.2 kDa, including the ha tag), whereas the smaller (~52 kDa) Nta1p-ha is initiated at the (+31) ATG codon (predicted Nta1p-ha of 52.2 kDa, including the ha tag) (Fig. 3*D*). The results of mRNA mapping are consistent with this interpretation, inasmuch as the set of *NTA1* mRNAs contains both the species whose 5′ ends encompass the (+1) ATG and the species whose 5′ ends are located between the (+1) and the (+31) ATG (Fig. 3*D*). The 29-residue region between Asp-4 and Asp-34 in the larger Nta1p (Fig. 3, *B* and *D*) resembles mitochondrial translocation signals (von Heijne, 1986). However, more extensive testing will be required to verify the conjecture that the larger Nta1p species might be a mitochondrial protein.

A Null nta1 Mutant and Biochemical Aspects of Nt-amidase—A deletion/disruption allele of *NTA1* (Fig. 3*A*) was used to produce the *nta1-Δ1* mutant (see "Experimental Procedures"). As expected from the phenotype of the original *nta1* mutants (Figs. 1*B* and 2*B*), Asn-βgal and Gln-βgal but not the other normally short-lived X-βgals were long-lived in the *nta1-Δ1* mutant ($t_{1/2}$ > 10 h) (data not shown), whereas they were short-lived in the congenic *NTA1* strain ($t_{1/2}$ of ~ 3 and 10 min, respectively (Varshavsky, 1992)). The normally long-lived X-βgals (bearing stabilizing N-terminal residues) remained long-lived in the *nta1-Δ1* mutant. These results supported the conjecture that *NTA1* encodes an amidase specific for N-terminal Asn and Gln. These data also indicated that Nta1p is the only such amidase in *S. cerevisiae*.

In a biochemical test, purified, ^{35}S-labeled Asn-dihydrofolate reductase or Gln-dihydrofolate reductase (dihydrofolate reductase-based N-end rule substrates (Bachmair and Varshavsky, 1989)) were incubated with extracts prepared from *E. coli* that either expressed or lacked Nta1p, and then fractionated by isoelectric focusing in a polyacrylamide gel. Isoelectric points of both substrates became more acidic after incubation with the Nta1p-containing *E. coli* extract but not after incubation with the control (Nta1p-lacking) extract. Moreover, the isoelectric point of Met-dihydrofolate reductase, an otherwise identical protein bearing a non-amide N-terminal residue, was not altered by incubation with either of *E. coli* extracts. These findings[2] confirmed the inferred deamidating activity of Nta1p and the confinement of this activity to N-terminal Asn and Gln.

Stewart *et al.* (1994) have purified a distinct Nt-amidase from porcine liver. They also isolated a cDNA encoding this enzyme (Stewart *et al.*, 1995). In contrast to the 52-kDa yeast Nta1p, which deamidates either N-terminal Asn or N-terminal Gln, the activity of the 33-kDa porcine Nt-amidase is confined to N-terminal Asn. This finding (Stewart *et al.*, 1994) suggests the existence of yet another mammalian Nt-amidase (the one specific for N-terminal Gln) and hence a bifurkation at the deamidation step in the N-end rule pathway of mammals but not of yeast. The amino acid sequence of the Asn-specific mouse Nt-amidase, deduced from sequences of the corresponding cDNA and the gene,[3] is highly similar to the sequence of porcine Nt-amidase but lacks similarities to the sequence of yeast Nta1p.

No significant phenotypic differences (other than the metabolic stabilization of Asn-βgal and Gln-βgal) were observed between the *nta1-Δ1* and congenic wt strains (see "Experimental Procedures"). Ubr1p (*N*-recognin) was recently found to be required for the peptide import in *S. cerevisiae*; *ubr1Δ* mutants do not express *PTR2*, which encodes a peptide transporter, and are unable to import peptides from the medium (Alagramam *et al.*, 1995). It is unknown whether this function of Ubr1p is mediated by the N-end rule pathway or another Ubr1p-dependent mechanism. Unlike *ubr1Δ* mutants, the *nta1-Δ1* mutant is able to import peptides.[4]

Overexpression of R-Transferase and Nt-amidase—Bartel *et*

[2] R. T. Baker, S. Grigoryev, and A. Varshavsky, unpublished data.
[3] S. Grigoryev, A. Stewart, S. Arfin, R. Bradshaw, and A. Varshavsky, unpublished data.
[4] C. Byrd and A. Varshavsky, unpublished data.

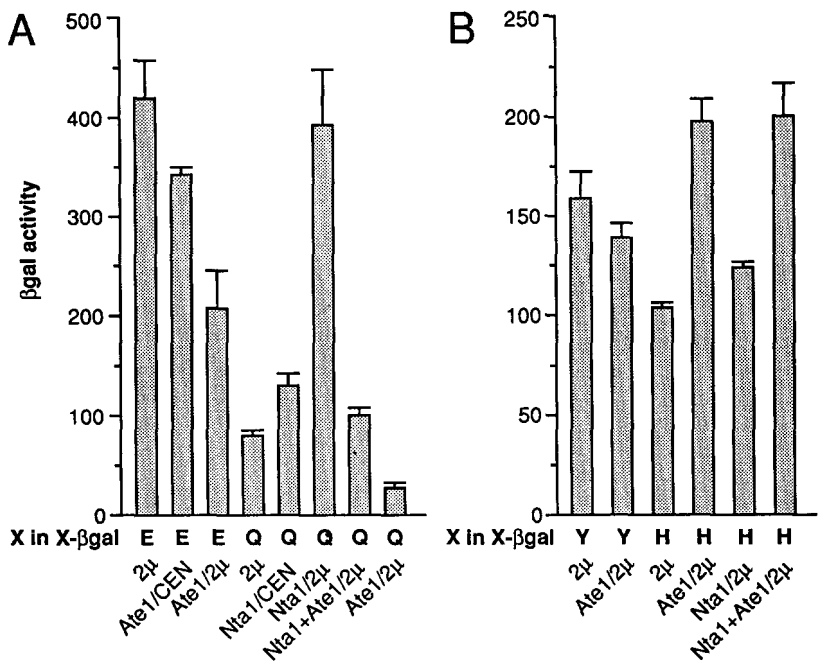


FIG. 6. **Effects of overexpressing Nt-amidase and R-transferase.** *A*, levels of βgal activity in extracts from cells overexpressing Ate1p (R-transferase) and/or Nta1p (Nt-amidase), and also expressing Glu-βgal (Ub-Glu-βgal) or Gln-βgal (Ub-Gln-βgal) (indicated by *E* and *Q*, single-letter abbreviations of Glu and Gln). The strain BWG1–7a expressing either Glu-βgal or Gln-βgal was transformed with either the high-copy vector YEplac195 (control) or an otherwise identical plasmid expressing either Ate1p or Nta1p from their natural promoters. Alternatively, BWG1–7a was transformed with a low copy (*CEN*-based) plasmid expressing either Ate1p or Nta1p from their natural promoters, as indicated. Values of βgal activity shown are the means of at least three independent measurements. Standard deviations are indicated *above* the *bars*. *B*, same as in *A* but with cells expressing either Tyr-βgal (Ub-Tyr-βgal) or His-βgal (Ub-His-βgal) (indicated by *Y* and *H*, single-letter abbreviations of Tyr and His).

al. (1990) found that overexpression of *S. cerevisiae N*-recognin (Ubr1p) accelerated the degradation of N-end rule substrates. We asked whether R-transferase (Ate1p) and Nt-amidase (Nta1p) are also rate limiting for certain classes of these substrates. Overexpression of R-transferase from a high copy plasmid in cells expressing either Glu-βgal or Gln-βgal decreased the levels of βgal activity by ~2- and ~3-fold, respectively (Fig. 6*A*). Thus, the arginylation of N-terminal Glu in Glu-βgal by R-transferase appears to be rate-limiting for the degradation of Glu-βgal and Gln-βgal. Even a weaker overexpression of R-transferase (from a low copy plasmid) resulted in a small but significant decrease of Glu-βgal (Fig. 6*A*). Previous work (Bartel *et al.*, 1990; Balzi *et al.*, 1990; Baker and Varshavsky, 1991; Dohmen *et al.*, 1991; Madura *et al.*, 1993) has shown that the concentration of an X-βgal test protein in yeast cells is a sensitive indicator of its metabolic stability.

Surprisingly, overexpression of Nt-amidase in cells expressing Gln-βgal *increased* the level of Gln-βgal by ~5-fold (Fig. 6*A*). In other words, overexpression of Nt-amidase *inhibited* the degradation of Gln-βgal. This result is likely to be related to an earlier finding that Gln-βgal, which bears an N-d^t residue and therefore requires two modifications (deamidation and arginylation) prior to its binding by *N*-recognin, has a *shorter* half-life ($t_{1/2}$ of ~10 min at 30 °C) than Glu-βgal ($t_{1/2}$ of ~30 min), which bears an N-d^s residue and is therefore only one step (arginylation) away from its binding by *N*-recognin (Bachmair and Varshavsky, 1989; Gonda *et al.*, 1989). No such "inverse" order of half-lives was observed with Asn-βgal and Asp-βgal ($t_{1/2}$ of ~3 min for both substrates) (*op. cit.*). The following assumptions are sufficient to account for these apparently paradoxical findings: (i) R-transferase arginylates Asp-βgal significantly faster than Glu-βgal; (ii) Nt-amidase is about equally effective in deamidating Asn-βgal and Gln-βgal; (iii) in wt cells, Nt-amidase exists largely as an Nt-amidase·R-transferase complex. Specifically, the R-transferase-mediated arginylation of Glu-βgal that has been produced from Gln-βgal by the Nt-amidase·R-transferase complex is presumed to occur kinetically in preference to the arginylation of Glu-βgal that reaches this complex directly from the bulk solvent, a feature known as "substrate channeling" in other multistage enzymatic reactions (Srere, 1987; Ovádi, 1991; Knowles, 1991; Negrutskii and Deutscher, 1991; Knighton *et al.*, 1994).

This model accounts for the observed stabilization of Gln-βgal upon overexpression of Nt-amidase. Indeed, under these conditions, a greater fraction of Gln-βgal is converted into Glu-βgal by the free (overexpressed) Nt-amidase. As a result, a greater fraction of the deamidation-produced Glu-βgal will have to reach the Nt-amidase·R-transferase complex directly from the bulk solvent, a kinetically inefficient route to arginylation. The resulting delay in formation of Arg-Glu-βgal (which can be bound by *N*-recognin) would cause the observed stabilization of Gln-βgal in cells that overexpress Nt-amidase (Fig. 6*A*). Note that overexpression of Nt-amidase raised the level of Gln-βgal to that of Glu-βgal (Fig. 6*A*). This result is also predicted by the model, because the bulk of Gln-βgal in cells that overexpress Nt-amidase is deamidated by the free (overexpressed) Nt-amidase rather than by the less abundant form of Nt-amidase that exists in the complex with R-transferase. Another prediction of the model is that expression of R-transferase and Nt-amidase is likely to be coregulated in wt cells to maintain optimal ratios of these apparently interacting enzymes. This conjecture is consistent with the presence of common sequence motifs in the 5′ regions of genes that encode components of the N-end rule pathway (Fig. 7 and below).

Overexpression of R-transferase Perturbs the Function of N-Recognin: Evidence for a Targeting Complex—Overexpression of R-transferase accelerated the degradation of Gln-βgal and Glu-βgal, which bear, respectively, an N-d^t and an N-d^s residue (Fig. 6*A*). However, the same overexpression *inhibited* the degradation of N-end rule substrates bearing a type 1 N-d^p residue (Fig. 6*B*). The yeast *N*-recognin and its mammalian counterparts contain a binding site for type 1 (basic) N-d^p residues Arg, Lys, and His, and another physically distinct binding site for type 2 (bulky hydrophobic) N-d^p residues Phe, Leu, Trp, Tyr, and Ile (Reiss *et al.*, 1988; Gonda *et al.*, 1989; Baker and Varshavsky, 1991; Varshavsky, 1992). Overexpression of R-transferase increased by ~2-fold the steady state level of His-βgal (bearing a type 1 N-d^p residue), and slightly but reproducibly decreased the level of Tyr-βgal (bearing a type 2 N-d^p residue) (Fig. 6*B*). Overexpression of Nt-amidase resulted in a slight inhibition of His-βgal degradation, whereas overexpression of both Nt-amidase and R-transferase caused a stronger inhibition of His-βgal degradation equal to the one observed upon overexpression of R-transferase alone (Fig. 6*B*). Thus,

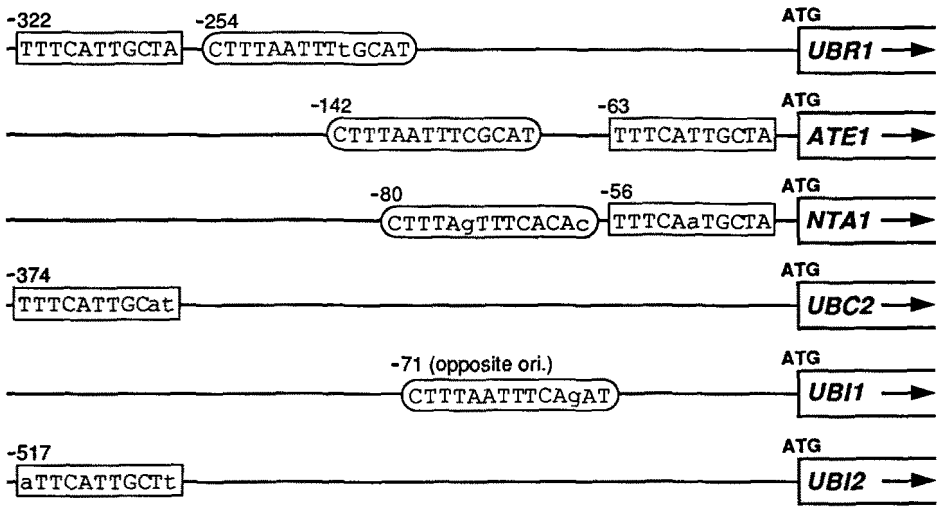


FIG. 7. **Common sequence motifs in promoters of genes encoding components of the N-end rule pathway.** Alignment of the 5′ regions of the *S. cerevisiae* genes *UBR1* (Bartel *et al.*, 1990), *ATE1* (Balzi *et al.*, 1990), and *NTA1* (the present work; Fig. 3, *A* and *D*) revealed two distinct regions of similarity: an 11-mer motif 1 (consensus TTTCATTGCTA) and a 14-mer motif 2 (consensus CTTTAATTTC*R*CAT; *R* = purine). Mismatches to the consensus are in *lowercase*. *Arrows* show the direction of transcription. While no other *S. cerevisiae* gene in GenBank contains both motifs, several genes, including three that encode known components of the Ub system, contain one of these motifs. Motifs 1 or 2 in the genes of this class (*UBC2*, *UBI1*, and *UBI2*) (Dohmen *et al.*, 1991; Özkaynak *et al.*, 1987) are also shown. The numbers indicate locations of the motifs relative to either the known (*UBC2*, *UBI1*, *UBI2*) or inferred start codons.

overexpression of R-transferase interferes with the function of the type 1 binding site in *N*-recognin, but slightly stimulates its type 2 binding site. The latter effect is consistent with the data indicating that an occupation of the type 1 site in either yeast or mammalian *N*-recognins with dipeptides bearing a type 1 N-d^p residue stimulates the activity of the other (type 2) site in *N*-recognin (Gonda *et al.*, 1989; Baker and Varshavsky, 1991).

These findings (Fig. 6*B*) are especially illuminating in conjunction with the data suggesting the existence of an Nt-amidase·R-transferase complex (Fig. 6*A*). Taken together, our results suggest that the 58-kDa R-transferase is physically associated with the 225-kDa *N*-recognin in proximity to its type 1 binding site. The "proximity" aspect of the postulated complex is invoked to account for the markedly different effects of overexpressed R-transferase on the functions of type 1 and type 2 binding sites in *N*-recognin (Fig. 6*B*). Specifically, a physical proximity of the bound R-transferase to the type 1 site in *N*-recognin is presumed to decrease the steric accessibility of this site to an N-end rule substrate bearing a type 1 N-d^p residue that approaches the type 1 site directly from the bulk solvent. Conversely, a substrate that acquired Arg (a type 1 N-d^p residue) through the arginylation by *N*-recognin-bound R-transferase would get access to the (nearby) type 1 binding site of *N*-recognin in kinetic preference to an otherwise identical substrate that has to reach the type 1 site directly from the bulk solvent.

This model (Fig. 8) postulates an Nt-amidase·R-transferase·*N*-recognin complex in which the access to the type 1 binding site of *N*-recognin directly from the bulk solvent may be partially obstructed by the bound Nt-amidase·R-transferase complex. Similarly, the access to the active site of R-transferase from the bulk solvent is presumed to be at least partially obstructed by the bound Nt-amidase. In this view, which accounts for the entire set of otherwise paradoxical interference data in Fig. 6, the spatially distinct type 2 binding site of *N*-recognin would not be inhibited by the presence of the Nt-amidase·R-transferase complex near the type 1 binding site of *N*-recognin, thereby explaining the observed dichotomy between the effects of overexpressed R-transferase on the reactions mediated by the type 1 and type 2 binding sites of *N*-recognin (Fig. 6*B*). The postulated targeting complex in Fig. 8 includes the Ubc2p Ub-conjugating enzyme, whose physical association with *N*-recognin was demonstrated directly (Dohmen *et al.*, 1991; Madura *et al.*, 1993).

N-recognin may partition *in vivo* between R-transferase-bound and free states. In this view, the free *N*-recognin would bind substrates bearing either type 1 or type 2 N-d^p residues directly from the bulk solvent, whereas the R-transferase-bound form of *N*-recognin would be preferentially accessible to substrates bearing N-d^t, N-d^s, or type 2 N-d^p residues (in comparison to substrates bearing type 1 N-d^p residues). Substrates bearing N-d^t or N-d^s residues would be "channeled" to the type 1 binding site of *N*-recognin after their modification by the *N*-recognin-bound Nt-amidase·R-transferase complex (Fig. 8). The mechanics of channeling may involve diffusion of an N-end rule substrate in proximity to surfaces of the targeting complex, similar to the channeling mechanism described for the bifunctional enzyme dihydrofolate reductase-thymidylate synthetase, where the channeling of dihydrofolate results from its movement across the surface of the protein (Knighton *et al.*, 1994). Overexpression of R-transferase would partition more of *N*-recognin into an R-transferase-bound form that is less active toward substrates bearing type 1 N-d^p residues, resulting in a slower degradation of these substrates, as observed (Fig. 6*B*).

Common Sequence Motifs in Promoters of Genes That Encode Components of the N-end Rule Pathway—We examined 5′ regions of *NTA1*, *ATE1*, and *UBR1* for common sequence elements and found two of them, 11- and 14-bp long, at different distances from the (inferred) start codons in each of these loci (Fig. 7). The 11-bp sequence TTTCATTGCTA (motif 1) is present in both *UBR1* and *ATE1*; a single-mismatch variant of this sequence is also present in *NTA1*. Variants of the 14-bp consensus sequence CTTTAATTTCRCAT (motif 2) are also present in *NTA1*, *ATE1*, and *UBR1* (Fig. 7). No other *S. cerevisiae* gene in data bases contains both of these motifs, suggesting that these sequences are recognized by transcriptional regulators whose combination is specific for genes encoding targeting components of the N-end rule pathway. The 5′ regions of about 15 *S. cerevisiae* genes in data bases contain one or the other but not both of the two motifs. Three of these genes, *UBI1*, *UBI2*, and *UBC2*, encode components of the Ub system. *UBI1* and *UBI2* encode identical precursors of Ub (Özkaynak *et al.*, 1987).

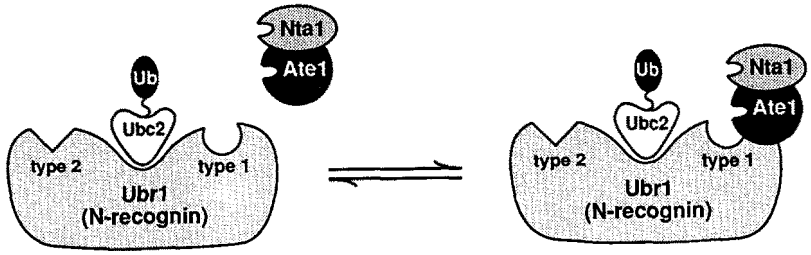


FIG. 8. **Model of a targeting complex in the *S. cerevisiae* N-end rule pathway.** The Ub-conjugating enzyme Ubc2p, which has been shown to be physically associated with Ubr1p (*N*-recognin) (Dohmen *et al.*, 1991; Madura *et al.*, 1993), is depicted carrying activated Ub linked to Cys-88 of Ubc2 through a thioester bond (Sung *et al.*, 1990). In the diagram, Ate1p (R-transferase) reversibly binds to Ubr1p in proximity to the type 1 substrate-binding site of Ubr1p. Furthermore, Nta1p (Nt-amidase) reversibly binds to Ate1p (or possibly to both Ate1p and Ubr1p; not shown). While this model accounts for the apparently paradoxical effects of overexpressed Ate1p (R-transferase) and/or Nta1p (Nt-amidase) on the activity of the N-end rule pathway (see the main text), the postulated targeting complex remains to be demonstrated directly.

UBC2 (Fig. 7) encodes the Ubc2p Ub-conjugating (*E2*) enzyme that is physically associated with the *UBR1*-encoded *N*-recognin (Dohmen *et al.*, 1991; Madura *et al.*, 1993) (Fig. 8). However, unlike Ubr1p, Ate1p, and Nta1p, which appear to have no functions outside of the N-end rule pathway, Ubc2p has other functions as well, mediated by complexes of Ubc2p with recognins distinct from *N*-recognin (Sung *et al.*, 1990; Sharon *et al.*, 1991; Ellison *et al.*, 1991).

Concluding Remarks—The *NTA1*-encoded N-terminal amidase (Nt-amidase) of the yeast *S. cerevisiae* mediates the conversion of tertiary (Asn or Gln) into secondary (Asp or Glu) destabilizing N-terminal residues in a substrate of the N-end rule pathway. The hierarchical organization of N-end rule, with its tertiary ($N\text{-}d^t$), secondary ($N\text{-}d^s$) and primary ($N\text{-}d^p$) destabilizing residues, is a feature that is more conserved in evolution than the Ub dependence of N-end rule pathways or the precise identity of enzymatic reactions that mediate the hierarchy of destabilizing amino acids in an N-end rule. For example, in bacteria such as *E. coli*, which lack Ub and Ub-specific enzymes, the N-end rule has both $N\text{-}d^s$ and $N\text{-}d^p$ residues (it lacks $N\text{-}d^t$ residues) (Tobias *et al.*, 1991). However, the identities of $N\text{-}d^s$ residues in *E. coli* (Arg and Lys) are different from those in eukaryotes (Asp and Glu in *S. cerevisiae*, Asp, Glu, and Cys in rabbit reticulocytes) (Varshavsky, 1992). Bacterial and eukaryotic enzymes that implement the coupling between $N\text{-}d^s$ and $N\text{-}d^p$ residues are also different: Leu,Phe-tRNA-protein transferase in *E. coli* and R-transferase in eukaryotes (Shrader *et al.*, 1993; Balzi *et al.*, 1990; Ciechanover *et al.*, 1988). Nonetheless, both bacterial Leu,Phe-tRNA-protein transferase and eukaryotic R-transferase catalyze reactions of the same type (conjugation of an amino acid to an N-terminal residue of a polypeptide) and use the same source of activated amino acid (aminoacyl-tRNA).

Hierarchical organization of the N-end rule "distributes" domains that recognize specific destabilizing N-terminal residues among several proteins such as Nt-amidase, R-transferase, and *N*-recognin. It is likely that cells can produce different *N*-recognins and can also regulate either synthesis or activity of Nt-amidase and R-transferase. The resulting changes of N-end rule may occur in response to physiologically relevant alterations in the state of a cell, for example, during cell differentiation. A change of the N-end rule may provide a way to destroy a set of previously long-lived proteins or to stabilize a set of previously short-lived proteins. A variety of indirect evidence supports this conjecture (Varshavsky, 1992) but a definitive test remains to be done. Physiological substrates of Nt-amidase and R-transferase remain to be identified as well.

Experiments in which Nt-amidase, R-transferase, or both of these enzymes were overexpressed in *S. cerevisiae* suggested a substrate channeling in the N-end rule pathway and a specific organization of its multienzyme targeting complex (Figs. 6 and 8 and above). These ideas are consistent with the presence of two distinct sequence motifs in the promoter regions of genes encoding Nt-amidase, R-transferase, and *N*-recognin (Fig. 7). No other *S. cerevisiae* gene in data bases contains both of these motifs, suggesting that they are recognized by transcriptional regulators whose combination is specific for genes encoding targeting components of the N-end rule pathway.

Acknowledgments—We thank S. Grigoryev for his advice about the *in vitro* deamidation assay and B. Bartel, M. Hochstrasser, and K. Madura for helpful discussions.

REFERENCES

Alagramam, K., Naider, F., and Becker, J. M. (1995) *Mol. Microbiol.* **15,** 225–234
Ambler, R. P., Auffret, A. D., and Clarke, P. H. (1987) *FEBS Lett.* **215,** 285–290
Arnason, T., and Ellison, M. J. (1994) *Mol. Cell. Biol.* **14,** 7876–7883
Ausubel, F. M., Brent, R., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1992) *Current Protocols in Molecular Biology*, Wiley-Interscience, New York
Bachmair, A., and Varshavsky, A. (1989) *Cell* **56,** 1019–1032
Bachmair, A., Finley, D., and Varshavsky, A. (1986) *Science* **234,** 179–186
Baker, R. T. (1991) *Nucleic Acids Res.* **19,** 1945
Baker, R. T., and Varshavsky, A. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 1090–1094
Baker, R. T., Tobias, J. W., and Varshavsky, A. (1992) *J. Biol. Chem.* **267,** 23364–23375
Balzi, E., Choder, M., Chen, W., Varshavsky, A., and Goffeau, A. (1990) *J. Biol. Chem.* **265,** 7464–7471
Bartel, B., Wünning, I., and Varshavsky, A. (1990) *EMBO J.* **9,** 3179–3189
Boeke, J. D., LaCroute, F., and Fink, G. R. (1984) *Mol. & Gen. Genet.* **197,** 345–346
Chau, V., Tobias, J. W., Bachmair, A., Marriott, D., Ecker, D., Gonda, D. K., and Varshavsky, A. (1989) *Science* **243,** 1576–1583
Ciechanover, A. (1994) *Cell* **79,** 13–21
Ciechanover, A., Ferber, S., Ganoth, D., Elias, S., Hershko, A., and Arfin, S. (1988) *J. Biol. Chem.* **263,** 11155–11167
deGroot, R. J., Rümenapf, T., Kuhn, R. J., Strauss, E. G., and Strauss, J. H. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 8967–8971
Deveraux, J., Haeberl, P., and Smithies, O. (1984) *Nucleic Acids Res.* **12,** 387–395
Dohmen, R. J., Madura, M., Bartel, B., and Varshavsky, A. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 7351–7355
Dohmen, R. J., Wu, P., and Varshavsky, A. (1994) *Science* **263,** 1273–1276
Ellison, K. S., Gwozd, T., Prendergast, J. A., Paterson, M. C., and Ellison, M. J. (1991) *J. Biol. Chem.* **266,** 24116–24120
Field, J., Nikawa, J. I., Broek, D., MacDonald, B., Rodgers, L., Wilson, I. A., Lerner, R. A., and Wigler, M. (1988) *Mol. Cell. Biol.* **8,** 2159–2165
Finley, D., Özkaynak, E., and Varshavsky, A. (1987) *Cell* **48,** 1035–1046
Gietz, R. D., and Sugino, A. (1988) *Gene (Amst.)* **74,** 527–534
Gonda, D. K., Bachmair, A., Wünning, I., Tobias, J. W., Lane, W. S., and Varshavsky, A. (1989) *J. Biol. Chem.* **264,** 16700–16712
Guarente, L., Yocum, R. R., and Clifford, R. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79,** 7410–7414
Guthrie, C., and Fink, G. R. (eds) (1991) *Guide to Yeast Genetics and Molecular Biology*, Academic Press, New York
Hershko, A., and Ciechanover, A. (1992) *Annu. Rev. Biochem.* **61,** 761–807
Hill, C. P., Johnston, N. L., and Cohen, R. E. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 4136–4140
Hinnebusch, A. G., and Liebman, S. W. (1991) in *The Molecular Biology of the Yeast Saccharomyces* (Broach, J. R., Pringle, J. R., and Jones, E. W., eds) Vol. 1, pp. 627–735, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
Hochstrasser, M. (1992) *Curr. Opin. Cell Biol.* **4,** 1024–1031
Hoffman, C. S., and Winston, F. (1987) *Gene (Amst.)* **57,** 267–272
Ito, H., Fukuda, Y., Murata, K., and Kimura, A. (1983) *J. Bacteriol.* **153,** 163–168
Jentsch, S. (1992) *Annu. Rev. Genet.* **26,** 179–207
Johnson, E. S., Gonda, D. K., and Varshavsky, A. (1990) *Nature* **346,** 287–291
Johnson, E. S., Bartel, B., Seufert, W., and Varshavsky, A. (1992) *EMBO J.* **11,** 497–505
Knighton, D. R., Kan, C. C., Howland, E., Janson, C. A., Hostomska, Z., Welsh, K. M., and Mathews, D. A. (1994) *Nature Struct. Biol.* **1,** 186–194

Knowles, J. R. (1991) *J. Theor. Biol.* **152,** 53–55
Kozak, M. (1992) *Annu. Rev. Cell Biol.* **8,** 197–225
Madura, K., and Varshavsky, A. (1994) *Science* **265,** 1454–1458
Madura, K., Dohmen, R. J., and Varshavsky, A. (1993) *J. Biol. Chem.* **268,** 12046–12054
Negrutskii, B. S., and Deutscher, M. P. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 4991–4995
Nogi, Y., Yano, R., Dodd, J., Carles, C., and Nomura, M. (1993) *Mol. Cell. Biol.* **13,** 114–122
Nyunoya, H., and Lusty, C. J. (1984) *J. Biol. Chem.* **259,** 9790–9798
Ovádi, J. (1991) *J. Theor. Biol.* **152,** 1–22
Özkaynak, E., Finley, D., Solomon, M. J., and Varshavsky, A. (1987) *EMBO J.* **6,** 1429–1440
Reiss, Y., Kaim, D., and Hershko, A. (1988) *J. Biol. Chem.* **263,** 2693–2698
Rose, M. D., Novick, P., Thomas, J. H., Botstein, D., and Fink, G. R. (1987) *Gene (Amst.)* **60,** 237–243
Rothstein, R. (1991) *Methods Enzymol.* **194,** 281–301
Sharon, G., Raboy, B., Parag, H. A., Dimitrovsky, D., and Kulka, R. G. (1991) *J. Biol. Chem.* **266,** 15890–15894
Sharp, P. M., and Li, W.-H. (1987) *Nucleic Acids Res.* **15,** 1281–1295
Sherman, F., Fink, G. R., and Hicks, J. B. (1986) *Methods In Yeast Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
Shrader, T. E., Tobias, J. W., and Varshavsky, A. (1993) *J. Bacteriol.* **175,** 4364–4374
Sikorski, R. S., and Hieter, P. (1989) *Genetics* **122,** 19–27
Spence, J., Sadis, S., Haas, A. L., and Finley, D. (1995) *Mol. Cell. Biol.*, in press
Sprague, G. F., Jensen, R., and Herskowitz, I. (1983) *Cell* **32,** 409–415
Srere, P. A. (1987) *Annu. Rev. Biochem.* **56,** 89–124
Stewart, A. E., Arfin, S. M., and Bradshaw, R. A. (1994) *J. Biol. Chem.* **269,** 23509–23517
Stewart, A. E., Arfin, S. M., and Bradshaw, R. A. (1995) *J. Biol. Chem.*, in press
Sung, P., Prakash, S., and Prakash, L. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 2695–2699
Teem, J. L., and Rosbash, M. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80,** 4403–4407
Tobias, J. W., and Varshavsky, A. (1991) *J. Biol. Chem.* **266,** 12021–12028
Tobias, J. W., Shrader, T. E., Rocap, G., and Varshavsky, A. (1991) *Science* **254,** 1374–1377
Varshavsky, A. (1992) *Cell* **69,** 725–735
von Heijne, G. (1986) *EMBO J.* **5,** 1335–1342
Werner, M., Feller, A., Messenguy, F., and Piérard, A. (1987) *Cell* **49,** 805–813